ttH 3 Lepton Final State at 13TeV Analysis

Chao Wang

Supervisor: Cristinel Diaconu, Emmanuel Monnier, Lianliang Ma

Outline

- ttH 3 Lepton Final State at 13TeV Analysis
  - ttH Overview
  - Cut Based Study and Selection Optimization in Signal Region
  - Fake Estimation
- Conclusion

Top and Higgs: the Special Relationship

"Higgs Boson"

"Top Quark"

Yukawa Coupling

$$y_t = \sqrt{2} m_t / v \simeq 1$$

Special role for the top quark in electroweak symmetry breaking?

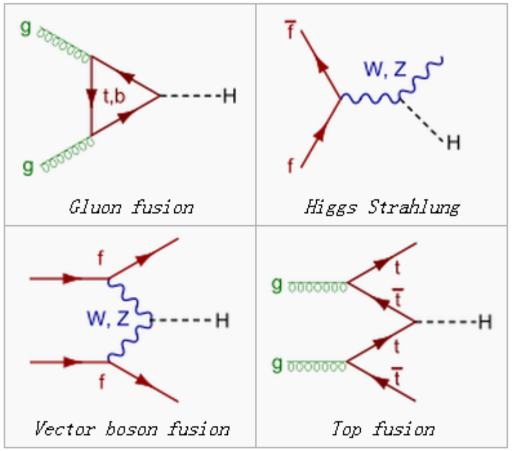
Close to unity - why?

Why is the top so heavy?

$t\bar{t}H$ production with smallest cross section.
Test properties and give direct access to top-Higgs coupling.

Higgs Boson

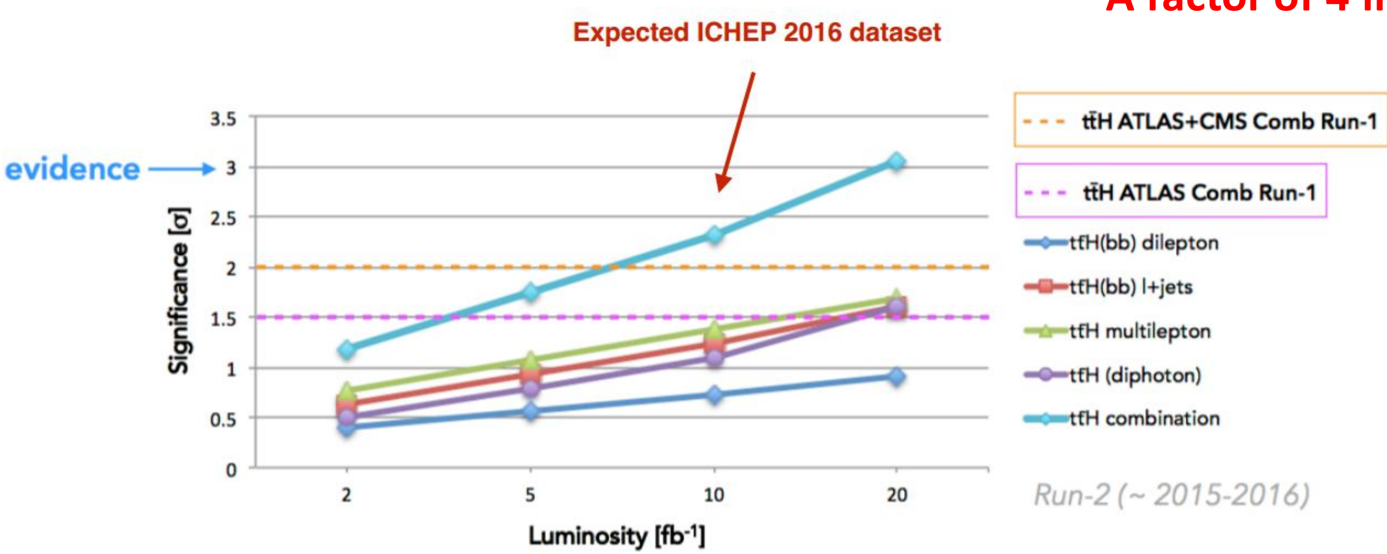
Higgs Boson (125GeV) Production at LHC



Cross section (pb) at $\sqrt{s} = 8$ (7) TeV	
ggF	19.52 (15.32)
VBF	1.58 (1.22)
WH	0.70 (0.57)
ZH	0.39 (0.31)
$t\bar{t}H$	0.13 (0.09)
Total	22.32 (17.51)

Cross Section(pb) at 13TeV	
ggF	43.92
VBF	3.748
WH	1.380
ZH	0.8696
ttH	0.5085

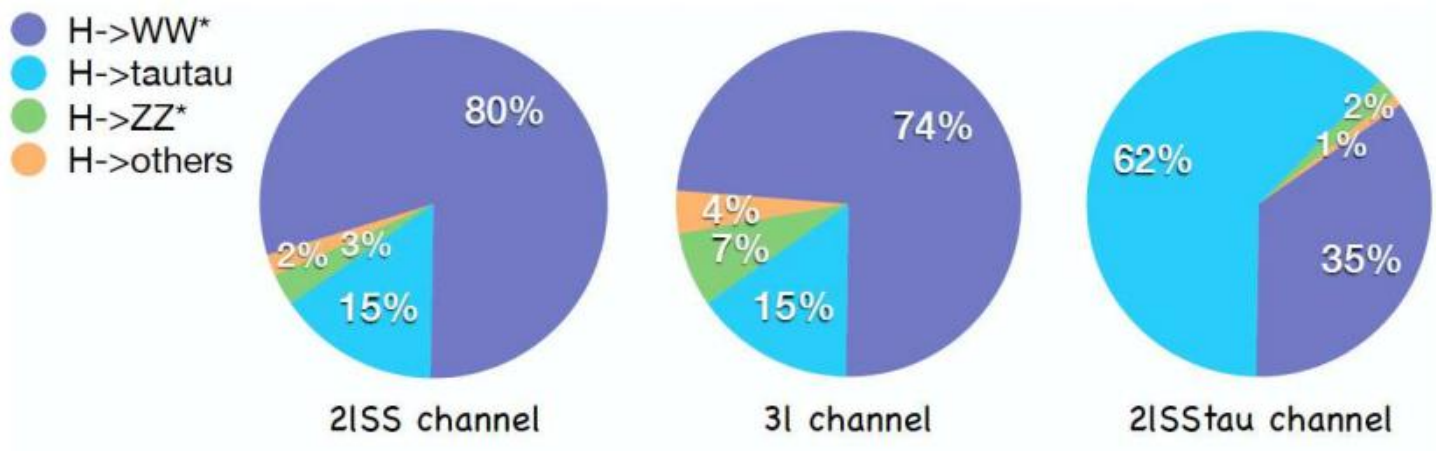
A factor of 4 in 13 TeV



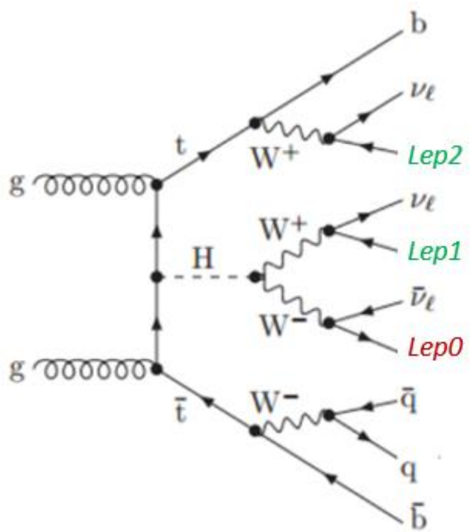
Run-2 (~ 2015-2016)

$t\bar{t}H$ Signature with Leptons

Signal



3l

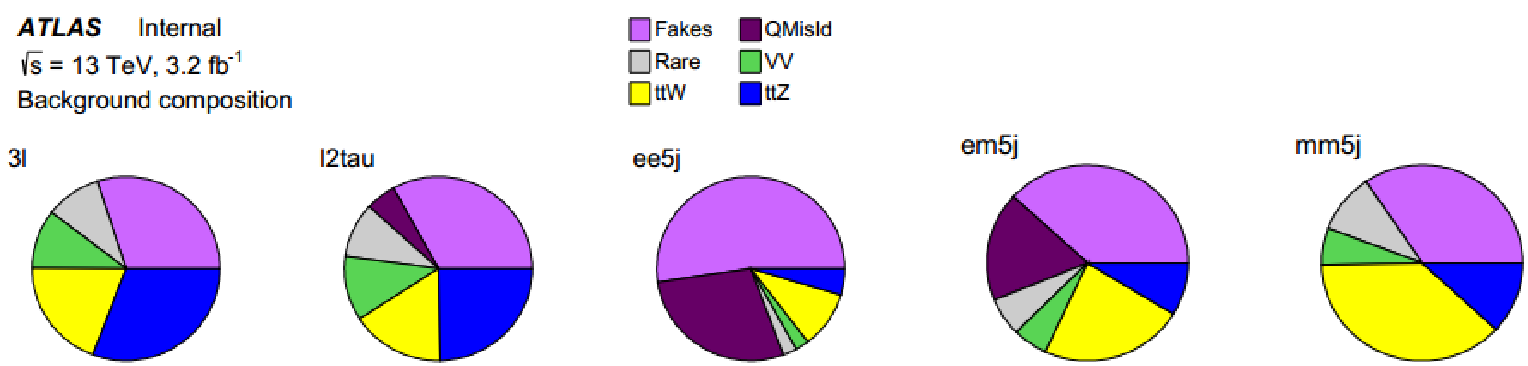


- **3 leptons, Lepton charges' sum = +/-1**
- **High jet multiplicity, b-jets**

Lep0: the lepton with no same-charge partner(never fake lepton for $t\bar{t}$, Z + jets background)
Lep1: ΔR closet to the Lep0

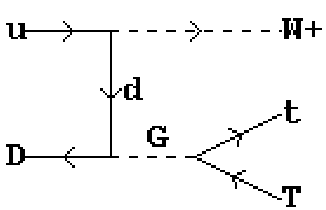
$t\bar{t}H$ Signature with 3 Leptons

Background

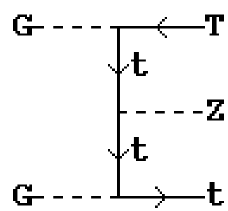


- Events same final state as the signal (irreducible backgrounds), three true charged leptons (MC):

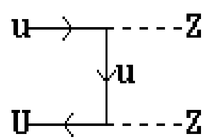
$t\bar{t}W$



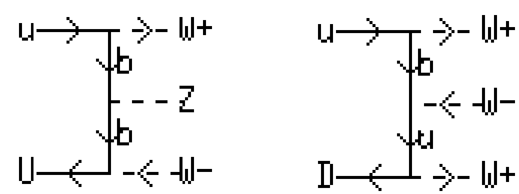
$t\bar{t}Z$



ZZ



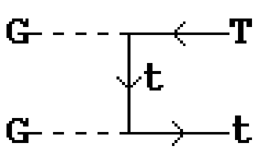
$W^{\pm}Z$



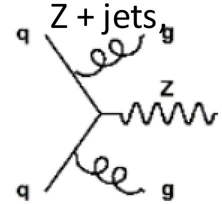
Obs: top-> bW ; W/Z can decay to leptons.

- Events with a non-prompt or a fake lepton selected as prompt lepton(DD):

$t\bar{t}$



Z + jets



Signal: leptons (some fake), low jet multiplicity, in general no b-jets.

Cut-based Study and Selection Optimization

Motivation

- Toward to Run2 Condition(higher energy scale and high pile-up)
- New b-tagging and some CP tools(Lep ID and Isolation…)

TTHtoLeptonsPreliminarySelection: "preliminary" selection for Moriond 2016 and Event Selection

https://twiki.cern.ch/twiki/bin/view/AtlasProtected/TTHtoLeptonsPreliminarySelection

Electrons

- pt > 10 GeV
- |eta| < 2.47, and not 1.37 < |eta| < 1.52
- pass LooseAndBLayer Likelihood ID
- |z0 sin theta| < 2 mm
- |d0 significance| < 10
- pass loose isolation

Tau Jets (hadronically decaying tau lepton)

- abs(charge==1)
- nTracks ==1 || nTracks ==3
- eta : [0, 1.37], [1.52, 2.5]
- JetIDBDTMedium == 1
- pT > 25 GeV
- EleOLR in TauSelectionTool

Muons

- pt > 10 GeV
- |eta| < 2.5
- pass loose muon quality
- z0 sin theta| < 2 mm
- |d0 significance| < 10
- pass loose isolation

Jets

- jet clean criteria ("LooseBad")
- pt > 25 GeV
- |eta| < 2.5
- |JVT|> 0.64 and |eta| < 2.4 and pt < 50 GeV

Selection in three leptons channel

- Require exactly 3 light leptons sum of lepton charge = +/- 1 :
- Require pt > 20 GeV for the two SS leptons (keep pt > 10 GeV for the OS lepton)
- Require any of the leptons matched to any of the triggers with 25 (21) GeV pT on the electron (muon)
- For OS, same-flavour pairs, require Z mass veto : [81, 101] GeV
- For OS, same-flavour pairs, require invariant mass > 12 GeV

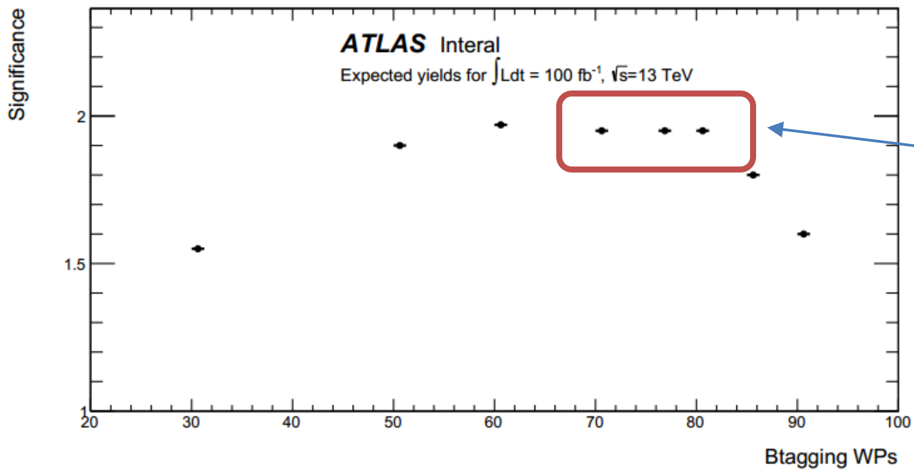
Selection Optimization -- b tagging and jet multiplicity

b tagging working point (MV2c20) and Jets multiplicity

All events are normalized to 100fb^-1

$$Z_0 \text{ significance estimator} = \sqrt{2((S+B)\ln(1+S/B)-S)}$$

Goal: test and find a good b tagging working point and Jet requirement

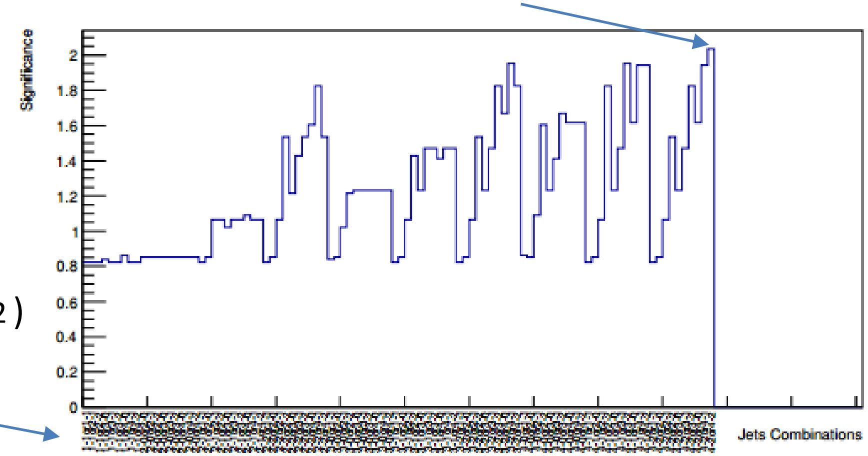


Flat performance from 70% - 80% WP
77% WP is chosen

(NJet >= 4 and NbJet >= 1) or (NJet >= 3 and NbJet >= 2)

Best jet combinations under 77% b-tagged efficiency.

Here 4-0 or 3-2 means: (NJet >= 4 and NbJet >= 0) or (NJet >= 3 and NbJet >= 2)



Significance vs jet combinations under 77% b-tagged efficiency

Selection Optimization -- Lepton ID and Isolation tool

Lepton ID and Isolation tool Study

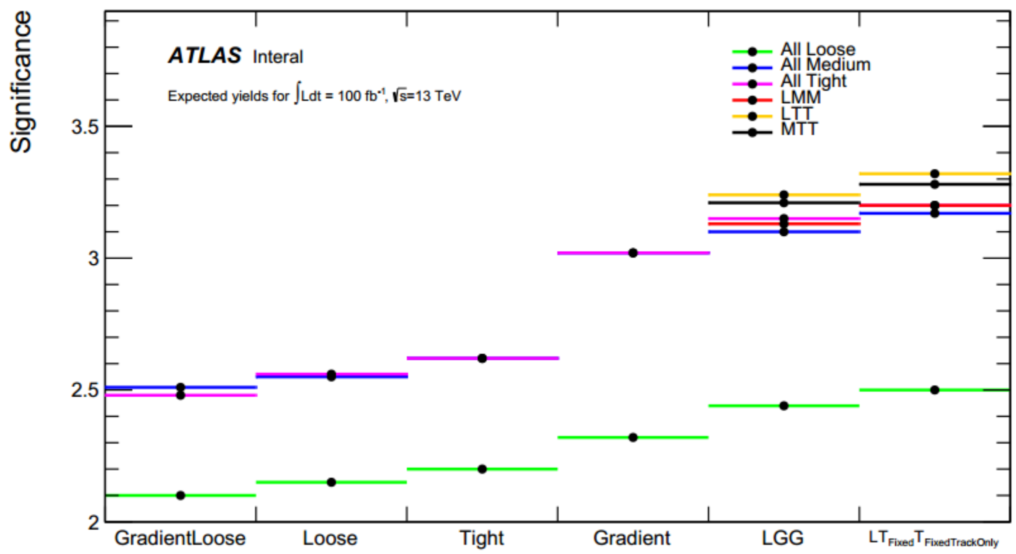
Goal: Isolation working point and lepton ID

- Lepton ID: loose, medium, tight

Jets' requirement:

4 jets of which at least one must be b-tagged or exactly 3 jets of which at least 2 are b-tagged

Different Isolation working points and lepton ID requirements are applied to **lep0, lep1, lep2**



LTT and $LT_{Fixed}T_{FixedTrackOnly}$:

lep0 is loose ID and loose isolation;

lep1, lep2 are tight ID and FixedCutTight(electron) FixedCutTightTrackOnly(muon) is the best

Fake Leptons

- Physics signatures with leptons in the final state is a large background rejection provided by the lepton identification

- Detector has limitations, like the spatial precision of the tracks or the efficiency to reconstruct leptons. A particle is found while it is not physically there. eg. if a charged particle other than a muon makes it to the MS, then it could be the case that the track it leaves in the MS, together with an arbitrary track in the ID will look like a muon to the algorithm's.

- Fake Origins:
  - Electrons fakes can arise from charged hadrons, photon conversions, or semi-leptonic heavy-flavor decays
  - Fake muons come from either semi-leptonic heavy flavor decays or meson decays in flight

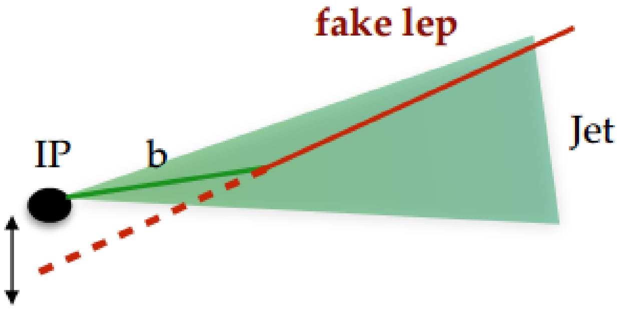
Expected and observed yields with Run1 Data(23fb^-1)

	Non-prompt	$t\bar{t}W$	$t\bar{t}Z$	Diboson	Expecxted bkg.	$t\bar{t}H$	Observed
3l	3.2 +/- 0.7	2.3 +/- 0.7	3.9 +/- 0.8	0.86 +/- 0.55	11.4 +/- 2.3	2.34 +/- 0.35	18

Non-prompt component is the dominating background (see Table)
A fake lepton is when a *b*-hadron decays semi-leptonically
To investigate fake leptons therefore requires very large samples of ttbar

Fake Lepton Estimation with ttbar

Fake leptons:
leptons produced inside jets decay, from heavy hadrons, photon conversion, light hadrons faking the electron shower, kaons or pions decays to mu



Partially **removed** **by:**
isolation requirements of surrounding tracks
Overlap removal:
removes jet close to e, remove u close to jets
impact parameter requirements

Method:
Analysis cuts on ‘tight’ leptons, additionally, define a ‘loose’ lepton with a relaxed isolation requirement or overlap removal

Transfer Factor(only with MC): **Tight /Anti-Tight = N(xll) /N(xl~~l(tight~~))**
Apply factor to the Data to estimate the fake lepton in the SR

$$N_{\text{estimated top},SR} = [N_{data,CR} - N_{\text{non}-topMC}] \cdot \theta$$

Fake Lepton Estimation with ttbar

Event Selection:

- ❑ 3l_charge = -/+1
- ❑ 3l_2 same sign pt > 20
- ❑ 3l_with Z veto opposite charge and same flavor
- ❑ invariant mass of opposite sign leptons > 12 GeV
- ❑ Jet requirement : jets with Pt>25 GeV and bjet(77%) (jet >=4 && bjet >=1) || (jet ==3 && bjet >=2)
- ❑ Z veto on mlll

Muon

TIGHT:

1) Loose ID
2) FixedCutTightTrackOnly Iso WP
3) |d0 significance| < 3
4) |z0 sin theta| < 0.5

⇩

Anti-Tight:

1) No overlap removal for muon
2) Isolation is reversed ptvarcone30/pt > 0.06 for muon

Electron

TIGHT:

1) Tight ID
2) FixedCutTight Iso WP
3) |d0 significance| < 5
4) |z0 sin theta| < 0.5 mm

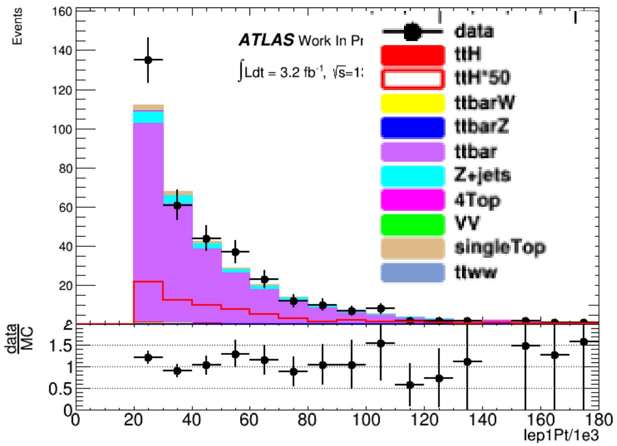
⇩

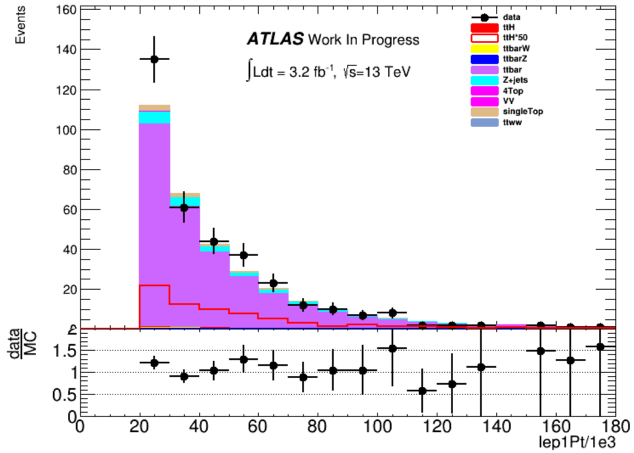
Anti-Tight:

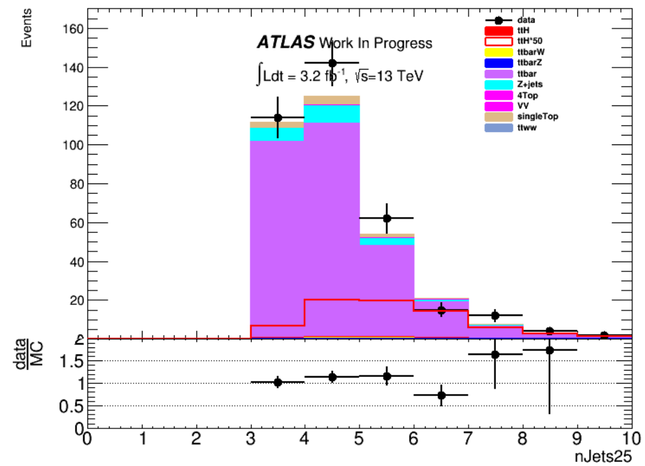
1) Loose or Medium ID (not Tight)
2) No isolation

CR Plots for Leptons

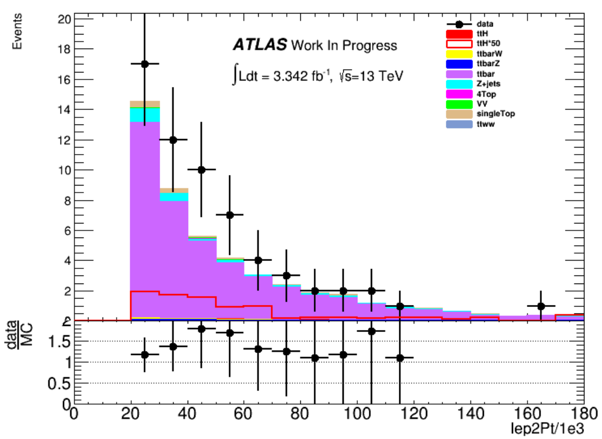
Muon (lep1pt, lep2pt, n_Jets)

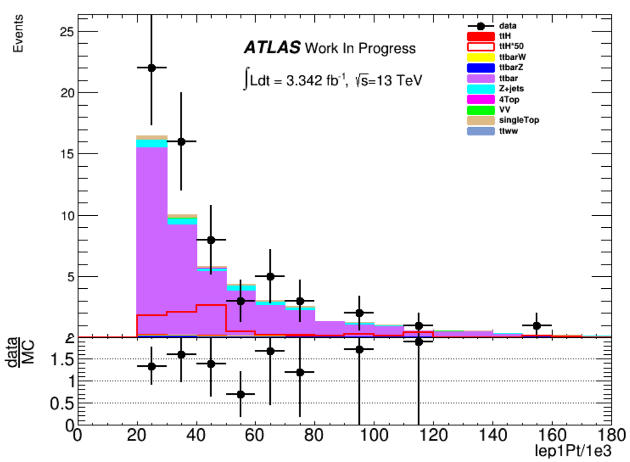


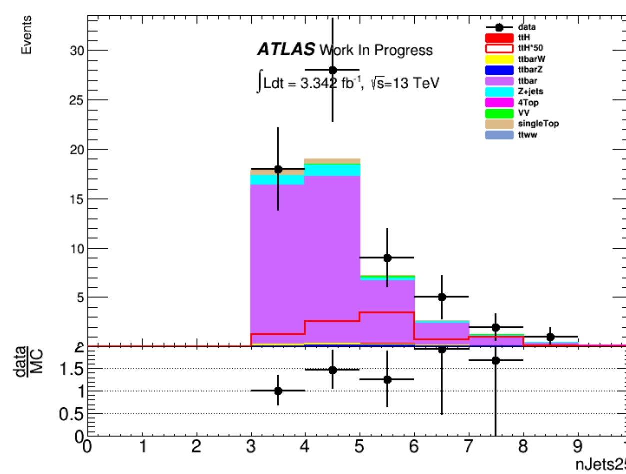




Electron (lep1pt, lep2pt, n_Jets)







Fake Leptons

Transfer Factor(only with MC): **Tight /Anti-Tight = N(xll) /N(xl~~l(tight~~))**

$$\theta_\mu = \frac{N(x\mu\mu\text{ events})}{N(x\mu p\text{ events})} \qquad \theta_e = \frac{N(xee\text{ events})}{N(xed\text{ events})}$$

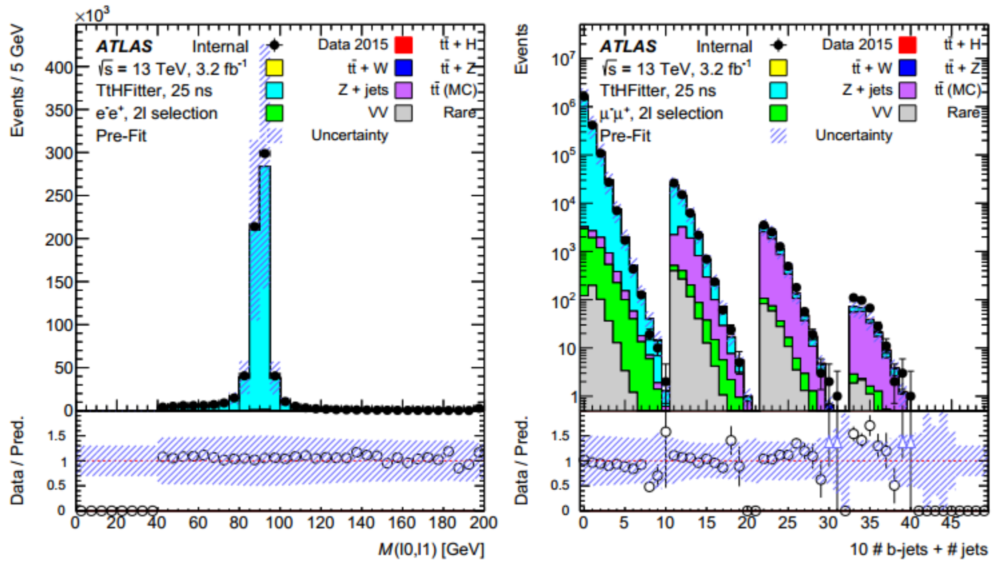
$$N_{\text{estimated top},SR} = [N_{data,CR} - N_{non-topMC}] \cdot \theta$$

	Sample	N in CR(Electron)	N in CR(Muon)
Control Region	Data	46 ± 6.8	315 ± 17.7
	$t\bar{t}$	30 ± 0.67	231.5 ± 1.9
	Others	0.83 ± 0.089	2.73 ± 0.20
Ratio	Data/MC	1.4	1.1
Signal Region	$t\bar{t}$	0.35 ± 0.071	0.206 ± 0.05
θ	N(xll)/N(xlp))	0.0261 ± 0.004	0.0015 ± 0.00034
Nest.top in SR		1.17 ± 0.34	0.47 ± 0.12

$\boxed{\text{Total Fake (muon+electron): 1.64 +/- 0.38}}$

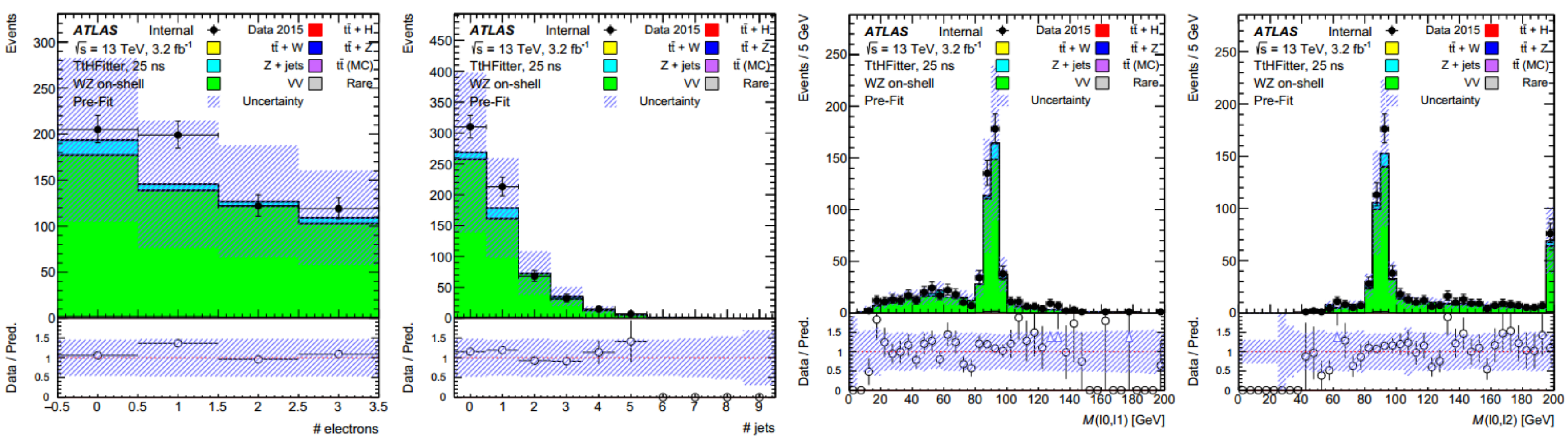
Validation Region

Z+jet Same Flavour



W±Z

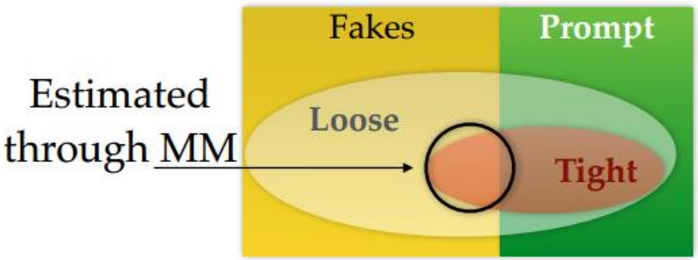
overall data and MC agreement is good



Matrix Method toward ICHEP

Why data driven: fakes are results of QCD processes, simulation not always reliable

- ✓ **Matrix Method gives differential background shapes, differently from cut and count, can be provided as input to MVA analysis, data driven technique to estimate background due to fake leptons**



4x4 Matrix in the 2 lepton case, *8x8* in the 3lep case

$$\begin{pmatrix} N_{tt} \\ N_{t\bar{t}} \\ N_{\bar{t}t} \\ N_{\bar{t}\bar{t}} \end{pmatrix} = M \begin{pmatrix} N^{ll}_{rr} \\ N^{ll}_{rf} \\ N^{ll}_{fr} \\ N^{ll}_{ff} \end{pmatrix}$$

Validation of the method on MC tt; Validation on data in the low jet multiplicity control region; Estimate of fakes in SR.

Matrix Method toward ICHEP

- Main limitation of Matrix Method for 2015 Data(3.2^fb-1):
  - too low statistics in fake control region if using isolation requirement on loose muons -> dominated by MC prompt and charge flip subtraction;
  - removing isolation requirement introduces two issues:
    self-consistency of the method: MM has to be applied on same type of “tight” and “loose” leptons on which measuring r and f;
    lepton counting on non-iso muons ≠ from analysis lepton counting;
    take into account OR with non isolated leptons;

Plans:

- ➢ Complete tests on loose lepton definition
- ➢ Evaluate trigger impact on r and f
- ➢ Provide MM results on 2016 early data on both 2lep and 3lep channels

Conclusion

- **Signal region Selection Optimization of ttH 3l final states**
  - b tagging working point, jet multiplicity, isolation working point and lepton ID
  - Baseline for 3l channel analysis

- **Estimation of fake background events(Theta Factor)**
  - As a cross check method for the fake estimation in 3l

- **Towards for ICHEP**
  - Matrix Method study will be tested for fake estimation in 3l

As analysis is still ongoing

Final expected yields, expected exclusion limits and signal strength are not shown


Thank you for your attention!

Backup

Cut-based Study and Selection Optimization

Sample Usage(MC15a samples):

- Z+jets: Madgraph+Pythia;
- ttW: MadGraphPythia;
- ttbar: PowhegPythia
- VV: Sherpa

ttH(semilep + dilep): aMcAtNloHerwigpp
ttZ: MadGraphPythia
ttbar(singletop, ttww, 4Top): PowhegPythia

• Lepton triggers

MC : HLT_mu20_iloose_L1MU15 | | HLT_mu50 | | HLT_e24_lhmedium_L1EM18VH | | HLT_e60_lhmedium | | HLT_e120_lhloose

Data : HLT_mu20_iloose_L1MU15 | | HLT_mu50 | | HLT_e24_lhmedium_L1EM20VH | | HLT_e60_lhmedium | | HLT_e120_lhloose

• Overlap removal(Run1)

Remove	**Overlap with**	**Cone size (Δ R)**
electron (low p_{T})	electron	0.1
electron	muon	0.1
jet	electon	0.3
muon	jet	$0.04 + 10/p_T^{\mu}$
tau	electron or muon	0.2
jet	tau	0.3

Validation Plots

VR	Selection
Inclusive OS dilepton (ee, $e\mu$, $\mu\mu$)	2 OS leptons, leading lepton $p_{\mathrm{T}} > 25$ GeV, sub-leading lepton $p_{\mathrm{T}} > 10$ GeV, $M(\ell^+\ell^-) > 40$ GeV, with 2ℓ lepton selections for muons and electrons Demonstrate normalization of Z, lepton scale factors
OS $t\bar{t}$ (top dilepton) (ee, $e\mu$, $\mu\mu$)	As for inclusive OS, but ≥ 2 jets and ≥ 1 MV2c20 77% b-tagged jets, ± 10 GeV veto around Z mass Demonstrate normalization of top, check btag scale factors
$t\bar{t}W$	2ℓ lepton selection require ≥ 2 b-tagged jets and ≤ 4 reconstructed jets H_T(jets) > 220 GeV in ee and $e\mu$ channel M_{ee} not within [75,105] GeV and $E_T^{miss} > 50$ GeV in ee channel
WZ on-shell	3ℓ lepton selection, require ≥ 1 OS SF pair within 10 GeV of Z and b-veto Verify WZ normalization, njet spectra
$W\ell\ell$ off-shell	As above but Z cut reversed Verifies off-shell modelling
WZ+HF on-shell	As for WZ on-shell but require ≥ 1 77% b tagged jets Verifies heavy flavor modelling for WZ
$t\bar{t}Z$	3ℓ lepton and jet selection require at least one OS SF pair within 10 GeV of $m_Z = 91.2$ GeV **+requiring 4j2b** Verify $t\bar{t}Z$ normalization, modelling

Fake Muon

Fake muon

- N(xuu): 17 0.206 +/- 0.05(weighted)
- N(xup): 10504 131.828 +/- 1.45(weighted)
  - theta factor: 0.0015 +/- 0.00034

$$\theta_\mu = \frac{N(x\mu\mu \text{ events})}{N(x\mu p \text{ events})}$$

(Only ttbar events are considered)

$$N_{\text{estimated top},SR} = [N_{data,CR} - N_{non-topMC}] \cdot \theta$$

315 +/- 17.7 events

2.732 +/- 0.202 events

0.47 +/- 0.12 events

Fake Electron

Fake electron

- N(xee): 29 0.347 +/- 0.069(weighted)
- N(xed)CR: 1185 13.27+/- 0.441(weighted)
  - theta factor: 0.026 +/- 0.0046

$$\theta_e = \frac{N(xee \text{ events})}{N(xed \text{ events})}$$

(Only ttbar events are considered)

$$N_{\text{estimated top},SR} = [N_{data,CR} - N_{\text{non}-topMC}] \cdot \theta$$

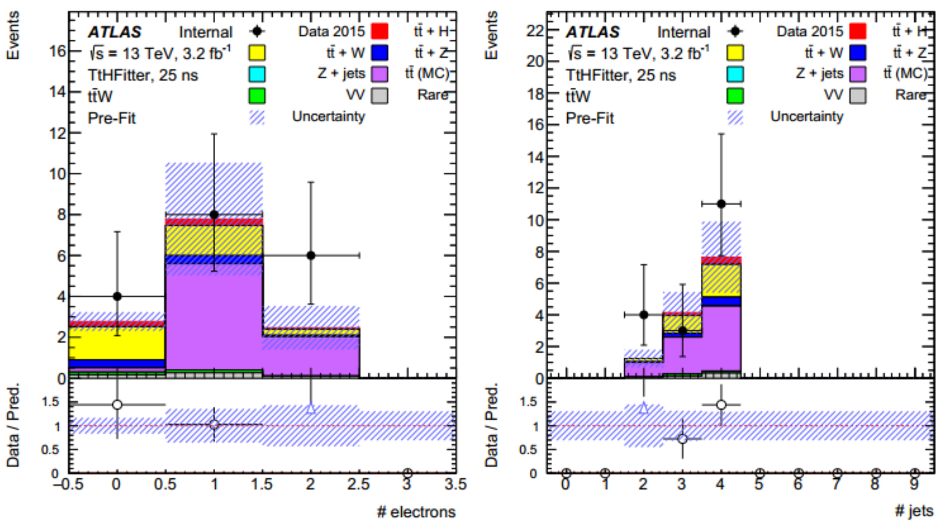
46 +/- 6.8 events

0.832 +/- 0.089 events

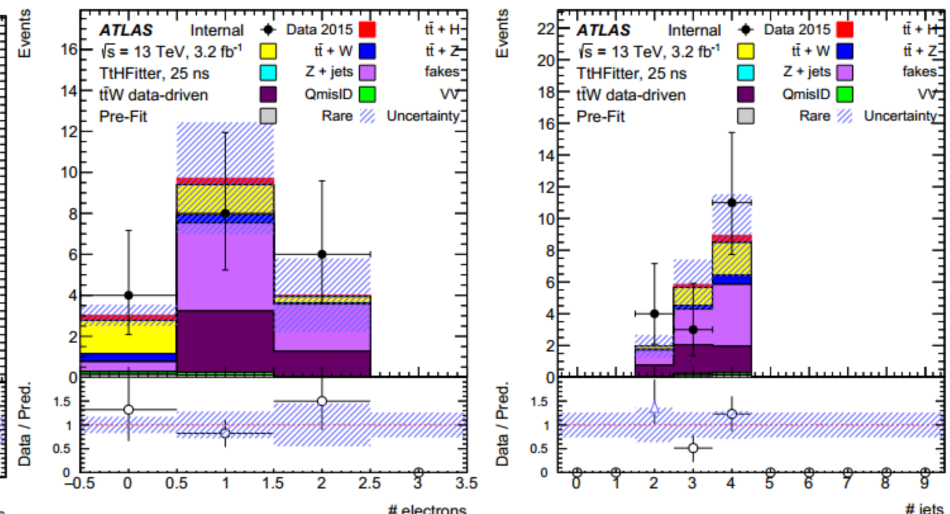
1.174 +/- 0.34 events

Total Fake (muon+electron): 1.64 +/- 0.38

ttW Validation Region



ttW validation region



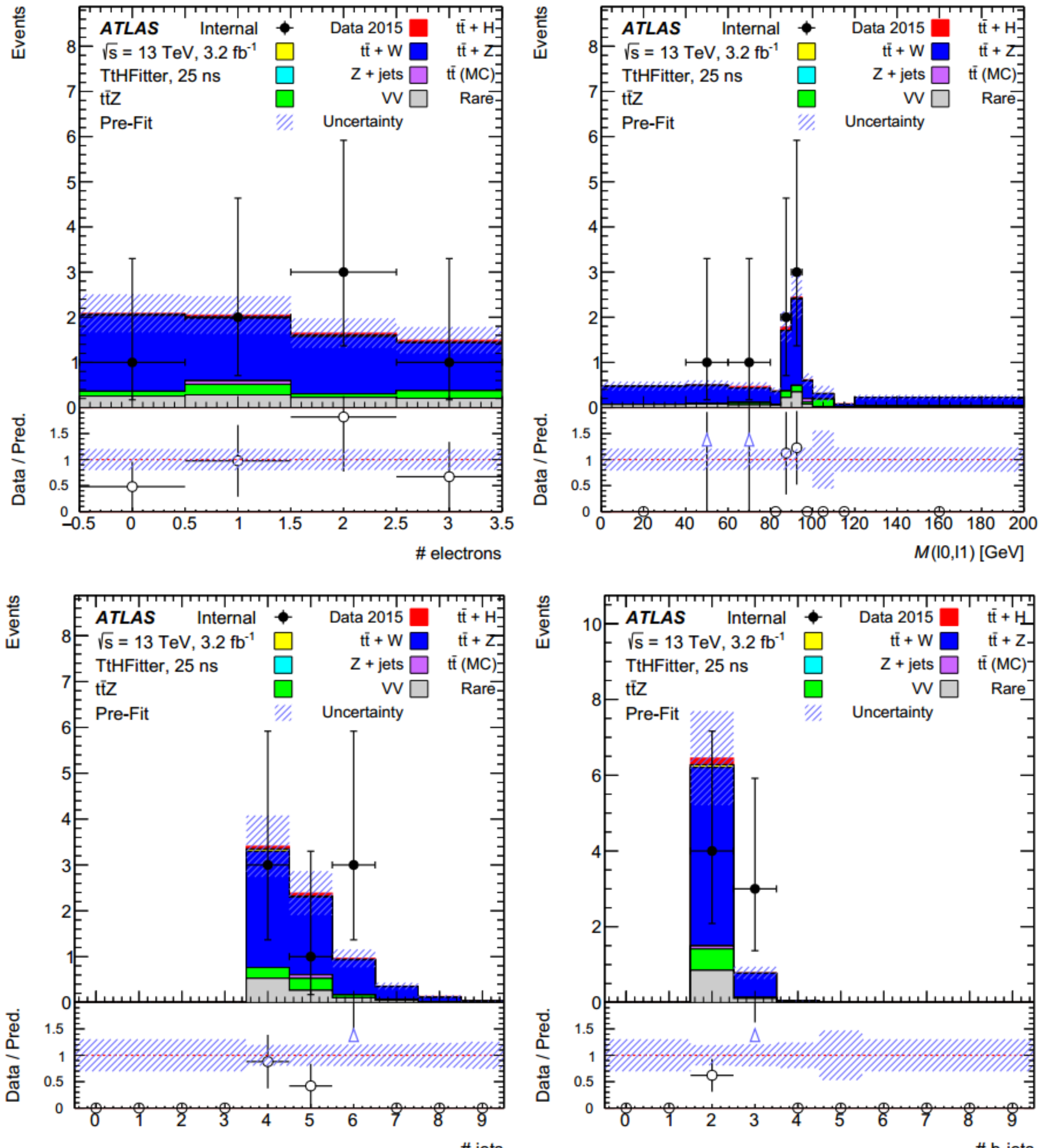
ttW validation region

data-driven QMisId and fakes estimates

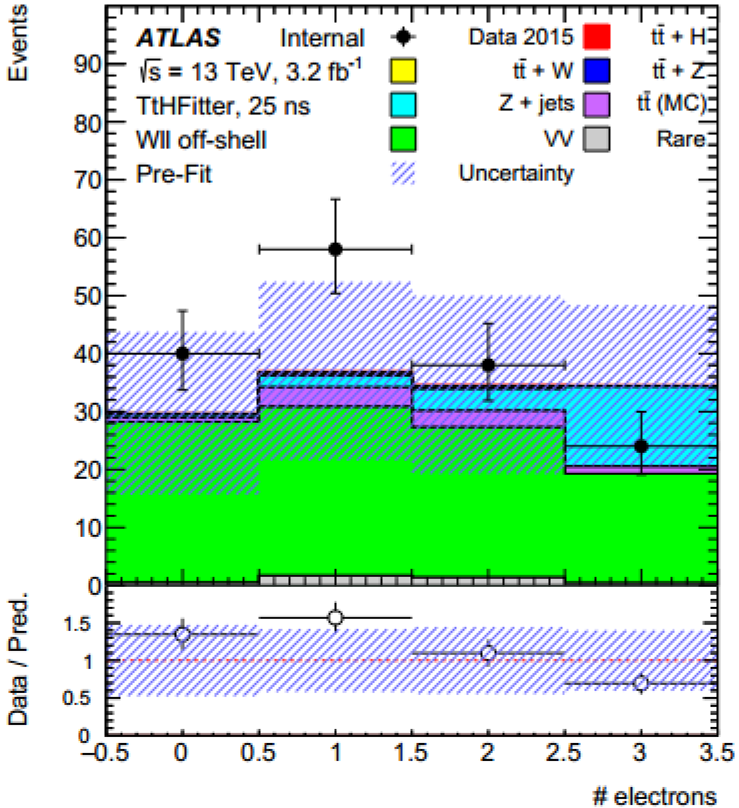
	$\mu^\pm\mu^\pm$	$e^\pm\mu^\pm$	$e^\pm e^\pm$
$t\bar{t}$ (MC)	0.205	5.20	1.92
fakes	0.49 ± 0.30	4.30 ± 2.47	2.30 ± 1.75
QMisId	0	2.99 ± 1.09	1.22 ± 0.38

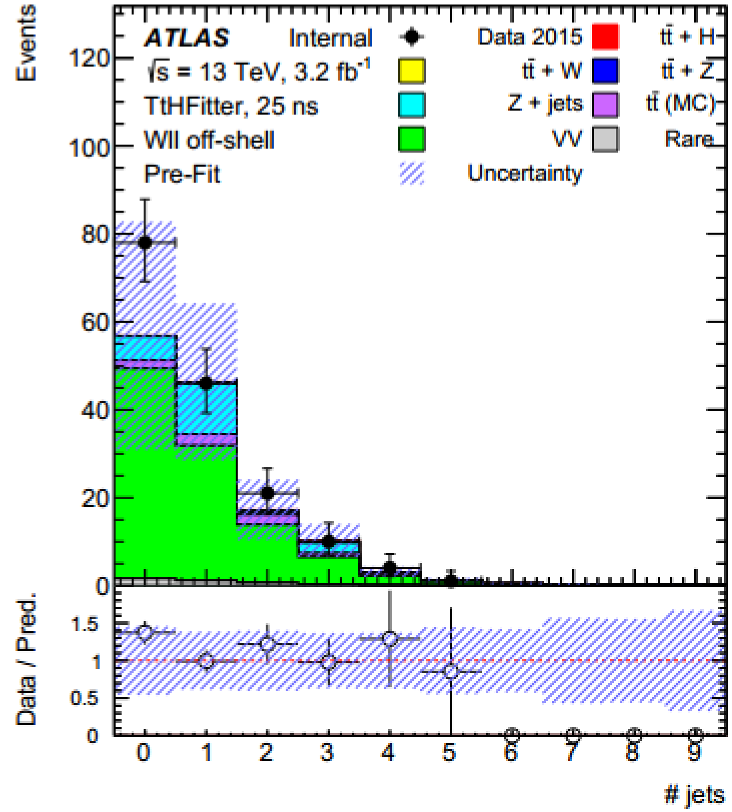
Yields for data-driven estimates of charge misidentification (QMisId) and fakes in different channels of $t\bar{t}W$ validation region.

ttZ Validation Region



Wll Validation Region





Run I and new recommended Overlap Removal

Run II.Shared tools and selections between ATLAS analysis aimed to satisfy three purposes:

- reduction of disk and CPU usage;
- reduction in the need to understand calibrate objects;
- harmonization of the analysis to ease comparison between groups.

The new recommended Overlap Removal procedure acts in the following way:

1. τ are removed when within a $\Delta R \leq 0.2$ from an electron or a muon.
2. Muons can be duplicated as electrons since they can radiate a hard photon. In this case the two objects are closer than 0.01 in ΔR or share the same ID track. The approach, when this duplication occurs, is to remove the electrons. The test in this section has been done requiring a $\Delta R \leq 0.01$ between electron and muon.
3. A jet is removed when closer in $\Delta R < 0.2$ to an electron, while the electron is later removed if closer to a jet in the $0.2 < \Delta R < 0.4$ region.
4. When a muon and a jet are overlapping, the jet is removed if closer than$\Delta R \leq 0.2$ and has less than at least 2 tracks in the inner detector.
5. The muon then is discarded when closer then 0.2 in Δ R distance. Eventually, a jet is removed when closer than $\Delta R < 0.2$ from a τ.

Significance: $S/\sqrt{S+B}$		
	ttHMultilep OR	ASG OR
2 leptons SS	0,1288	0.1269
3 leptons	0,1701	0,1665
2 leptons SS + τ	0,0877	0,0870

Run I overlap removal does not modify in a significant way in the *ttH* multilepton channel.
New recommended overlap removal scheme will be used in the subsequent version of this analysis

- Qualification task
    - Optimization of cut-based electron identification menus for Run2 at low Pt

Optimization of cut-based EID menus

Motivation:

- Lower energy background is higher due to low energy photo and pion
- Run2 has a higher luminosity than run 1 and pile-up situation us different
- Reconstruction and the identification menus have also to be retuned

  - Signal: $J/\psi \to ee$ and $Z \to ee$
  - Background: Inelastic minimum bias and dijet filtered, ET(jet) > 17 GeV
  - The 9 variables are considered due to their separation power are:

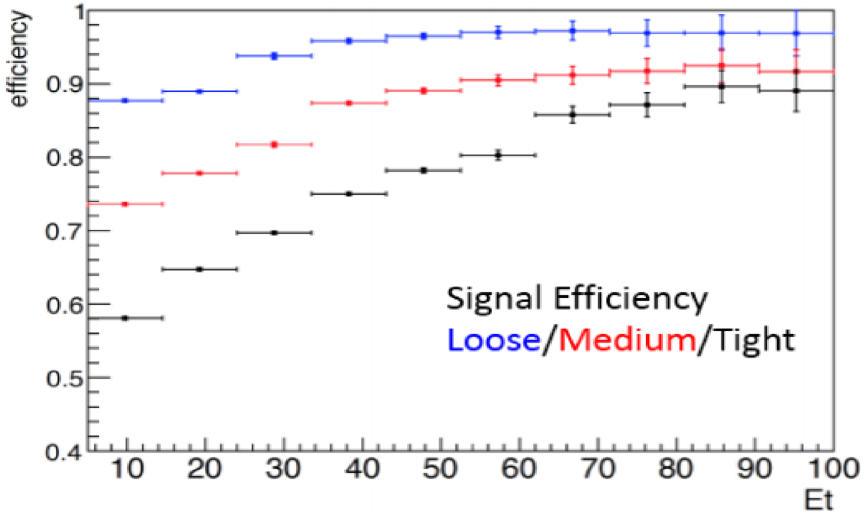
    f3, wstot, ERatio, RHad, $R\eta$, $R\phi$, $W\eta 2$, eOverP, $\Delta\Phi$

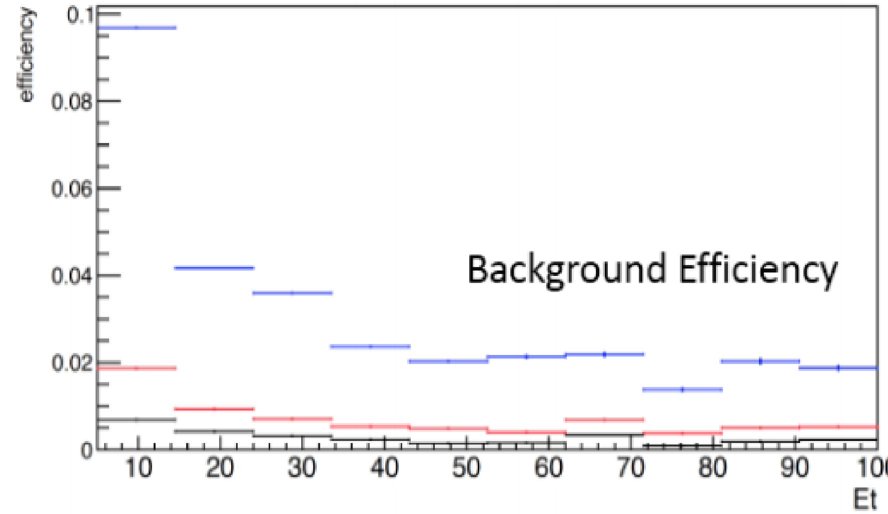
Strategy:

- Pre-selection
  - Truth match used for signal, and anti-truth match used for background – Et: [5,20] ; |eta|<2.47
- Cleaning cuts were applied to avoid cuts in tails of distributions
- Optimization carried out in groups of variables(correlation and pileup dependence) TMVA Cut-based method to optimize the cuts
- A target efficiency was decided for each low Et bin
  - Flatness in efficiency desired
- Post-optimization, cuts were smoothened
  - Avoid jumps in cut values from one bin to next
- Inclusiveness of menus ensured
  - Loose is looser than medium and medium is looser than tight
- Monotonicity of cut values as a function of Et
  - Cut value in the i-th eta bin always greater than or less than i-th eta bin in the next Et bin

Menus	5-10(GeV)	10-15(GeV)	15-20(GeV)
Loose	0.88	0.88	0.88
Medium	0.75	0.75	0.80
Tight	0.60	0.65	0.75

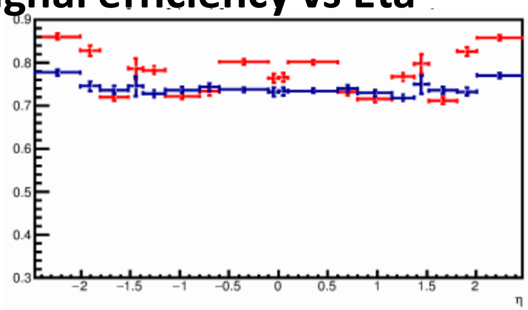
Optimization of cut-based EID menus

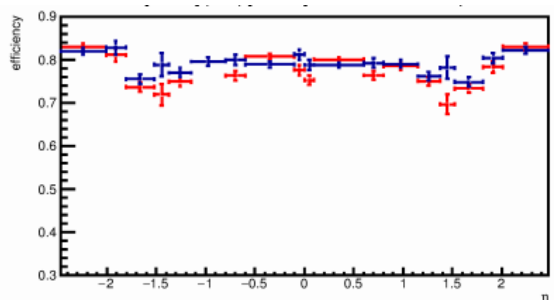


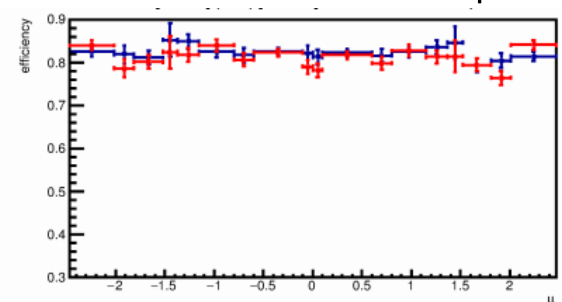


Thanks to Venu's plots

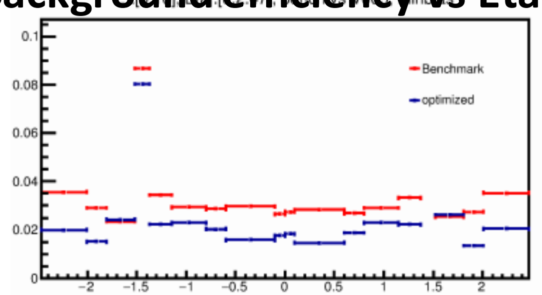
Signal efficiency vs Eta

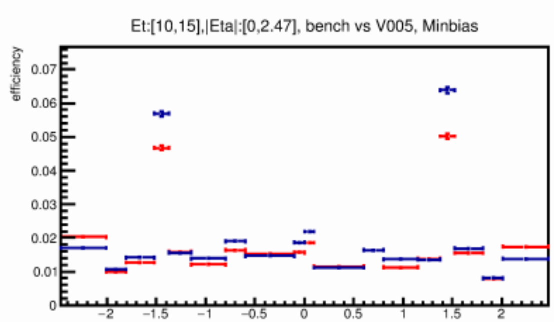


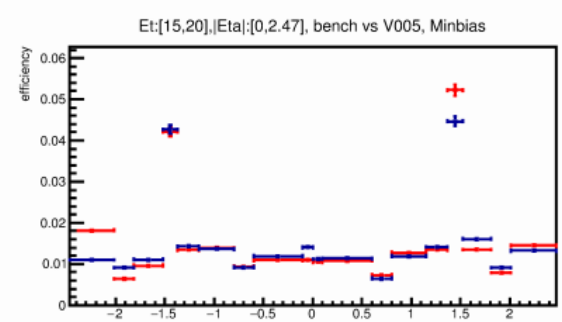




Background efficiency vs Eta







5-10 GeV **10-15 GeV** **15-20 GeV**